ORIGINAL RESEARCH

The Random Forest Model Has the Best Accuracy Among the Four Pressure Ulcer Prediction Models Using Machine Learning Algorithms

Jie Song [1]
Yuan Gao[2]
Pengbin Yin[3]
Yi Li[1]
Yang Li[2]
Jie Zhang[4]
Qingqing Su[1]
Xiaojie Fu[2]
Hongying Pi[5]

[1]Medical School of Chinese PLA, Beijing, People's Republic of China; [2]First Medical Center, Chinese PLA General Hospital, Beijing, People's Republic of China; [3]Fouth Medical Center, Chinese PLA General Hospital, Beijing, People's Republic of China; [4]Sixth Medical Center, Chinese PLA General Hospital, Beijing, People's Republic of China; [5]Medical Service Training Center, Chinese PLA General Hospital, Beijing, People's Republic of China

Correspondence: Hongying Pi
Medical Service Training Center, Chinese PLA General Hospital, No. 28 Fuxing Road, Haidian District, Beijing, 100853, People's Republic of China
Tel/Fax +86 010-66939159
Email pihongying@301hospital.com.cn

Purpose: Build machine learning models for predicting pressure ulcer nursing adverse event, and find an optimal model that predicts the occurrence of pressure ulcer accurately.

Patients and Methods: Retrospectively enrolled 5814 patients, of which 1673 suffer from pressure ulcer events. Support vector machine (SVM), decision tree (DT), random forest (RF) and artificial neural network (ANN) models were used to construct the pressure ulcer prediction models, respectively. A total of 19 variables are included, and the importance of screening variables is evaluated. Meanwhile, the performance of the prediction models is evaluated and compared.

Results: The experimental results show that the four pressure ulcer prediction models all achieve good performance. Also, the AUC values of the four models are all greater than 0.95. Besides, the comparison of the four models indicates that RF model achieves a higher accuracy for the prediction of pressure ulcer.

Conclusion: This research verifies the feasibility of developing a management system for predicting nursing adverse event based on big data technology and machine learning technology. The random forest and decision tree model are more suitable for constructing a pressure ulcer prediction model. This study provides a reference for future pressure ulcer risk warning based on big data.

Keywords: pressure ulcer, adverse event, machine learning, risk management

Introduction

Pressure ulcer, also known as pressure injury, refers to the localized injury of the skin and/or subcutaneous tissue, which usually occurs at the bone protuberance, and the part in contact with medical equipment or other equipment. It can be expressed as intact skin or open ulcers and may be accompanied with pain.[1] According to related research reports, the incidence of pressure ulcer in hospital is usually 2% ~5%, and the incidence of tape avulsion in elderly patients is as high as 15%.[2,3] Pressure ulcer can destroy the integrity of the skin, increase the risk of infection, and is difficult to heal. The high incidence, serious hazards and complex causes have made pressure ulcer a challenging issue, which attracts continuous attention in clinical care. In addition to the serious medical complications, people suffering from pressure ulcer are also faced with 229 kinds of physiological, social and psychological effects, which significantly affects the quality of life. The clinical intervention of pressure ulcer focuses on prevention. Research suggests that through dynamic monitoring and effective management, the occurrence of pressure ulcer

can be successfully prevented. Therefore, providing an effective early warning model of pressure ulcer to assist clinicians and nurses in making timely predictions and taking corresponding measures is of great value for the prevention of pressure ulcer.

It is identified that pressure ulcer is related to many risk factors, including continuous local pressure, hospital stay, long-term bed rest, neurological changes, etc.[4] Meanwhile, there are newly discovered factors every year. The Braden-Norton-Waterlow pressure ulcer assessment scale has been widely used in clinics and has made significant contributions to the management of pressure ulcer. However, the specificity and sensitivity of these scales are low, and there is still no evidence that these scales can effectively predict the occurrence of pressure ulcer. Recently, the big data and machine learning technology have undergone a fast-growing. These new technologies are able to directly extract data from the medical system for real-time analysis, and the accuracy of analysis is improved by huge amount of data.[5] In this case, these technologies are expected to solve many problems clinics, such as the pressure ulcer management.

This study aims to predict the pressure ulcer adverse events of inpatients through the machine learning technology. Firstly, the data of pressure ulcer in inpatients are analyzed and are associated with the electronic case data system. A logistic regression model is then used to identify 19 independent risk factors. These risk factors are input into four commonly machine learning algorithms to construct the prediction model. Finally, experimental results of the machine learning algorithms are investigated to find the algorithm with the best prediction performance.

Patients and Methods
Population
A total of 1839 patients are retrospectively included in the First Medical Center of the Chinese People's Liberation Army General Hospital during the period of hospitalization from January 1, 2013 to December 31, 2016. The included patients were over 18 years old and met the 2009 NPUAP pressure ulcer diagnostic criteria (the criteria adopted by the hospital). The false-positive data and suspicious cases as well as pressure ulcer events that occurred before hospitalization and within 24 hours after admission were excluded. In this study, the standard for defining false-positive cases is that after the nurse reports the patient's pressure ulcer event, the pressure ulcer

management team composed of three doctors contact the patient to diagnose the pressure ulcer, and give the final diagnosis result for pressure ulcer. If the result is consistent with that reported by the first-line nurse, the case is considered as pressure ulcer, otherwise, the case is considered as false positive. If the result cannot be defined, the case is considered as suspected. Finally, 1673 patients with pressure ulcer and 4141 patients without pressure ulcer were included.

Primary Outcome
The diagnostic definition of pressure ulcer in the adverse event reporting system of Chinese medicine hospital is based on the 2009 NPUAP Quick Reference Guide, where pressure ulcer refers to the local damage of the skin or/and subcutaneous tissue, usually located at the bone protrusion.[6] This kind of damage is generally caused by pressure or pressure combined with shear force. However, considering recent in-depth research on pressure ulcer, NPUAP has updated the definition of pressure ulcer. After comparing the diagnostic criteria of pressure ulcer in 2009 and 2019, all included cases meet the latest diagnostic criteria.

Variables
The predictor of pressure ulcer mainly includes general demographic data, basic vital signs data, medical care measures, disease-related data, and nursing evaluation items. Indicators are measured every day, all using the average value within 48 hours before the pressure ulcer occurs. In this study, the time of the pressure ulcer occurrence is randomly distributed. To match the time randomness of the pressure ulcer occurrence, the average data within 48 hours are randomly selected for analysis by computer (see Figure 1) during the period from 24 hours after admission to the discharge of patients in the control group.

The electronic medical system in the hospital records complete general demographic data, including gender, age (y), height (cm), and weight (kg). The basic vital signs review the data recorded in the "nursing workstation", which is developed by the hospital to inspect patients and performs nursing records. The inspection includes total intake (mL/day), total output (mL/day), body temperature (°C), systolic blood pressure (mmHg), and blood glucose (mmol/L).

In the electronic medical system, the medical care measures taken by the patient before the occurrence of
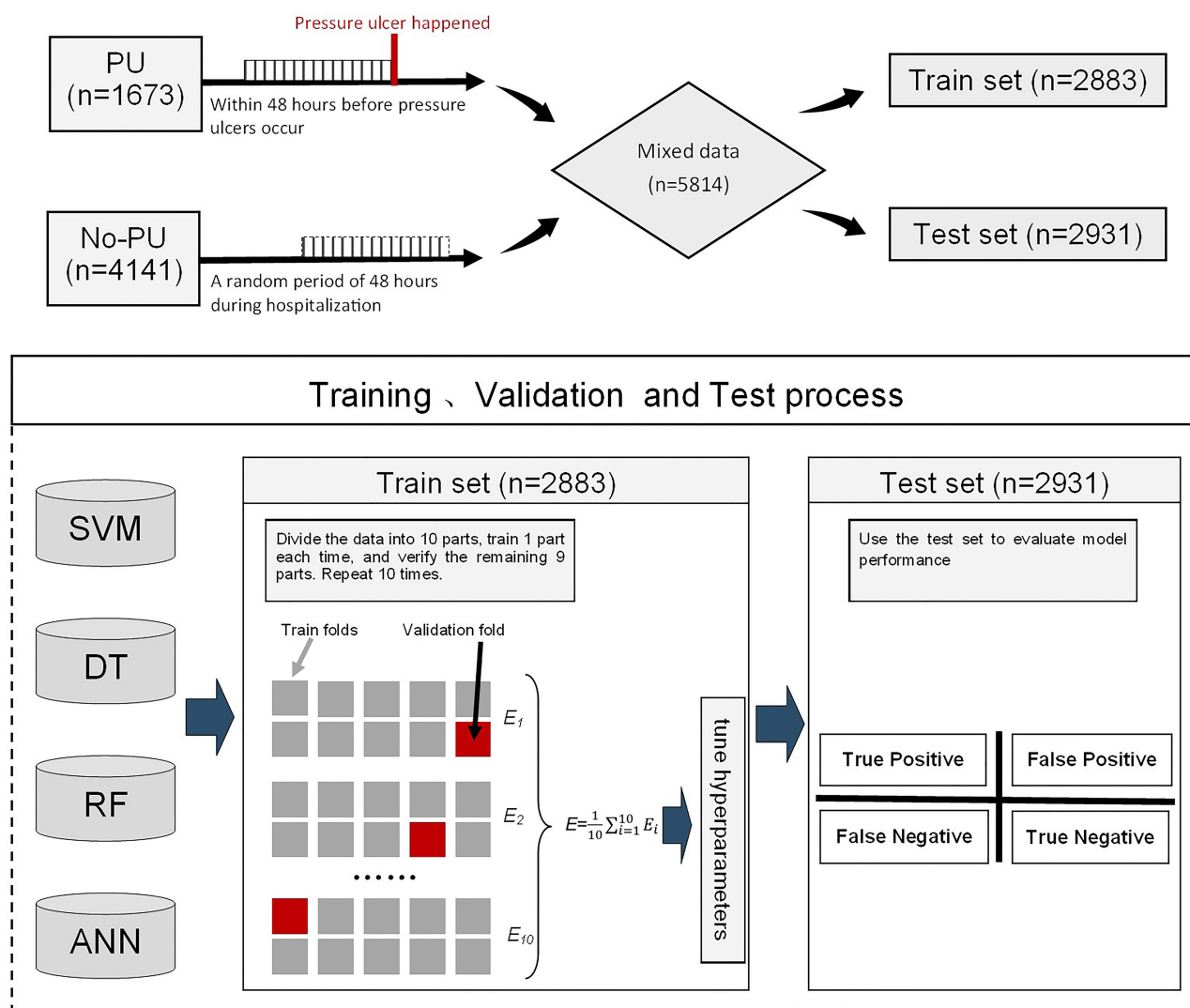
Figure 1 Flowchart of the model construction. The data of patients suffering from pressure ulcer was selected within 48 hours before pressure ulcer occurred. For patients without pressure ulcer, the data within 48 hours between 24 hours after admission and before discharge was randomly selected. Then, the two sets of data were fully mixed and randomly divided into two parts, namely train set (n=2883) and Test set (n=2931). Two sets. The model learns features in the train set, without knowing the actual pressure ulcer occurrence. Cross-validation was performed in the train set and the model performance was evaluated on the Test set.

pressure ulcer were found, including the length of stay (day), whether to stay in bed, whether to use restraint bands and whether to undergo surgery. These measures are very related to the pressure ulcer of hospitalized patients caused by medical equipment.

According to the disease diagnosis in the electronic medical record, we paid attention to whether the patient has diarrhea, diabetes, fractures and other related records. These variables are all related to the occurrence of pressure ulcer, and we incorporated these variables into the model for better accuracy.

Inpatients will undergo daily pressure ulcer assessments, passive turn-over status records and nutritional assessments. These data will be recorded in the medical

system in the form of electronic documents. The data is extracted and the average value of pressure ulcer within 48 hours before the occurrence of pressure ulcer is used. The evaluation of pressure ulcer in this study uses Norton pressure ulcer assessment (Norton scale). The sub-items of the evaluation scale consist of nutritional status score, mentality score, activity score, walking score, urinary and incontinence score. Further retrospective analysis found that some items have significance for predictions. Therefore, the incontinence score, activity score, mentality score, and the total score of the pressure ulcer assessment scale are used in the predictors (all scores). This study also counts the total scores of patients' nutritional risks, and the nutritional assessment follows the NRS2002 Nutritional

Risk Assessment Scale (Nutritional Assessment). Since position changes can reduce the risk of pressure ulcer, the patients that are assessed by clinical nurses to have risk of pressure ulcer will receive nursing services to change their positions every 2 hours, and the patients receiving care are recorded.

According to the documents and literature investigated by the hospital, 108 features were obtained from the hospital digital medical record database, including 5 general demographic data, 10 basic vital signs data, 34 medical care measures, 21 disease-related data, 8 nursing evaluation items, and 30 types of drugs. From these data, 19 features related to the occurrence of pressure ulcer were obtained through the logistic regression model.

Data Processing
Missing Value Filling
The data preparation in this study includes data elimination, missing value filling, as well as data format and unit unification, which was performed by two professionals good at biomedical information data processing. 7356 cases of data were extracted from the digital medical record database SQL server, including 1839 patients in the pressure ulcer group (PU) and 5517 patients in the non-pressure ulcer group (No-PU) (see Figure 2). To verify the accuracy of the data extracted from the digital medical record database, the value and timestamps in the extracted data were compared with that displayed in the clinician's electronic health record manually. As for implementation of the fully developed query for all manually validated cases, consistent values and timestamps (within 10 minutes) were obtained for all 7356 cases (100% agreement). Besides, the individual variables were cleaned by Stata 13 software (StataCorp LLC), and some non-conforming data were further excluded, including cases without data on the day of occurrence of pressure ulcer, repeated reported cases, unstructured data that cannot be processed and non-related data, as well as cases with more than 10% missing data. A total of 1184 unqualified data were eliminated, including 298 cases with pressure ulcer and 886 cases without pressure ulcer. In addition, the missing values were filled in the data. Specifically, on the day when the pressure ulcer event is reported, the average value of the corresponding structural data in the current natural day was taken; if there is no data on the day of reporting, the average value of the continuous data or the mode of non-continuous data in the previous three days was taken; if there no appropriate value can be

used, the case is eliminated. A total of 358 cases were eliminated in the second step. Also, 316 cases with missing values were filled and 6.3% of the data was repaired. Finally, the data format and unit were unified.

Data Segmentation
The data of 1673 patients with pressure ulcer event and 4141 patients without pressure ulcer event were mixed and merged. For machine learning algorithm using cross-validation as the evaluation method, the training data are split to realize a uniform sampling in the factor analysis.[7] Therefore, 50% of the patient data were randomly selected for training (n = 2883), and the other 50% were used for testing (n = 2931) (shown in Figure 1).

Model Fitting and Development
The construction of the model includes three stages: 1) Create a data set with no missing values; 2) Find 19 important variables through logistic regression; 3) Build a machine learning model based on the 19 variables.

The selection of important variables in stage 2 is fundamental. Though the predictive variables can be analyzed based on clinical experience, it is impossible to determine the variables that are actually important. Therefore, this study used a logistic regression model to verify the importance of the variables at the statistical level (selected $P<0.05$ variables), thus the best set of predictor variables can be found. By weighing the number of main variables, a balance is achieved between the model complexity (mainly dependent on the number of variables) and the model reliability.

Four methods are used to build prediction models, including support vector machine (SVM), decision tree (DT), random forest (RF) and artificial neural network (ANN) models. The performance of these four methods is compared. The DT-based model in this study uses the C5.0 algorithm with a minimum number of leaf nodes, which avoids the problem of too many branches in the ID3 algorithm. Also, pruning is performed during the construction of the decision tree to discretize continuous data, and the limit is set to the maximum number of leaf nodes. SVM-based model uses Gaussian inner product as the kernel function (SVM-Kernel). Through the iterative solution of sub-problems, the prediction of large-scale problems is finally completed. The gamma parameter in the model is set to 0.024. ANN is divided into an input layer, an output layer and a hidden layer. The information is collected through the input layer, and the data is input
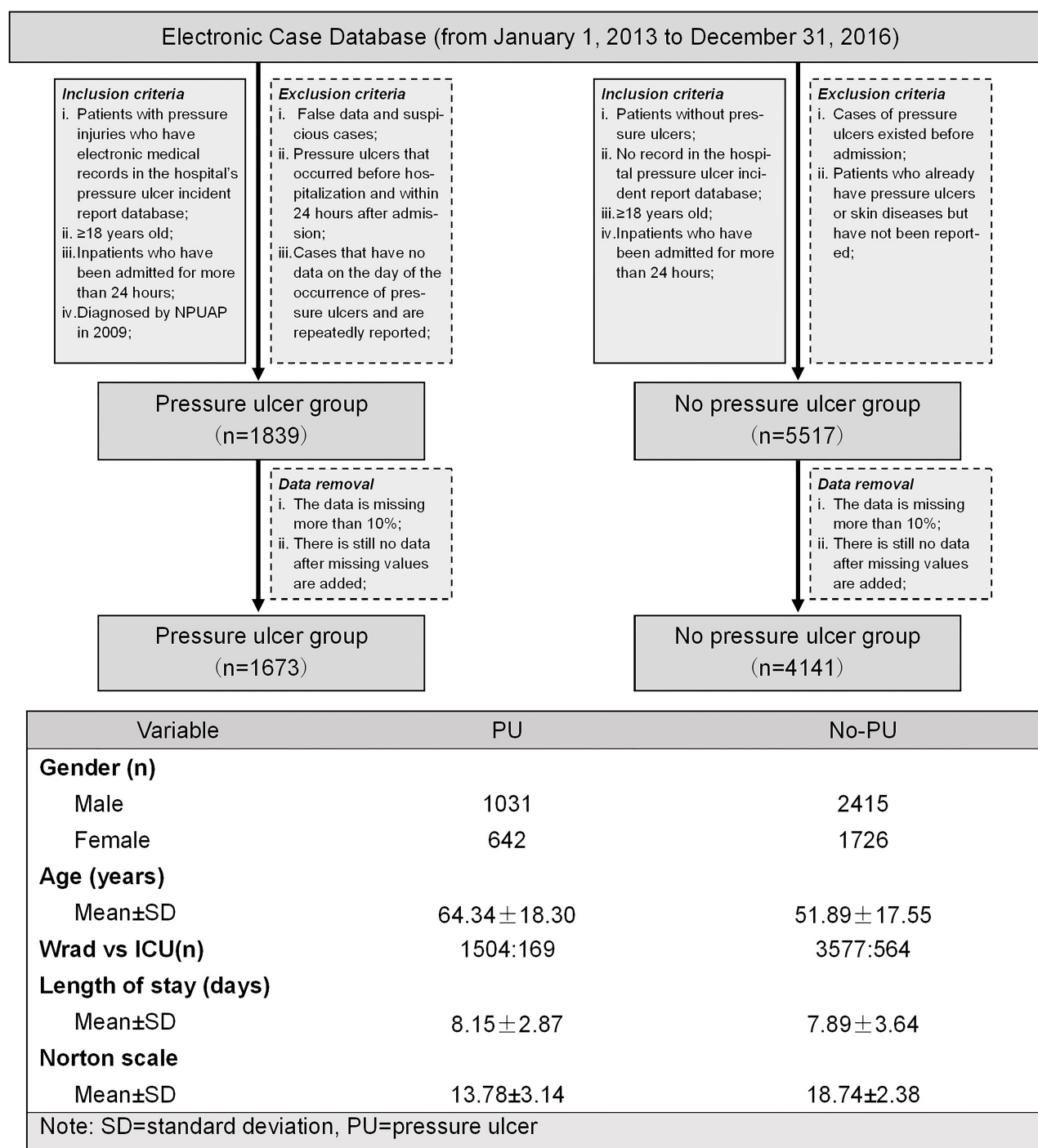
Electronic Case Database (from January 1, 2013 to December 31, 2016)

Inclusion criteria
i. Patients with pressure injuries who have electronic medical records in the hospital's pressure ulcer incident report database;
ii. ≥18 years old;
iii.Inpatients who have been admitted for more than 24 hours;
iv.Diagnosed by NPUAP in 2009;

Exclusion criteria
i. False data and suspicious cases;
ii. Pressure ulcers that occurred before hospitalization and within 24 hours after admission;
iii.Cases that have no data on the day of the occurrence of pressure ulcers and are repeatedly reported;

Inclusion criteria
i. Patients without pressure ulcers;
ii. No record in the hospital pressure ulcer incident report database;
iii.≥18 years old;
iv.Inpatients who have been admitted for more than 24 hours;

Exclusion criteria
i. Cases of pressure ulcers existed before admission;
ii. Patients who already have pressure ulcers or skin diseases but have not been reported;

Pressure ulcer group (n=1839)	No pressure ulcer group (n=5517)

Data removal
i. The data is missing more than 10%;
ii. There is still no data after missing values are added;

Data removal
i. The data is missing more than 10%;
ii. There is still no data after missing values are added;

Pressure ulcer group (n=1673)	No pressure ulcer group (n=4141)

Variable	PU	No-PU
Gender (n)		
Male	1031	2415
Female	642	1726
Age (years)		
Mean±SD	64.34±18.30	51.89±17.55
Wrad vs ICU(n)	1504:169	3577:564
Length of stay (days)		
Mean±SD	8.15±2.87	7.89±3.64
Norton scale		
Mean±SD	13.78±3.14	18.74±2.38
Note: SD=standard deviation, PU=pressure ulcer		

Figure 2 Flow diagram of data inclusion. The figure shows the data sources, data selection process, inclusion and exclusion criteria for patients with and without pressure ulcer.

into the hidden layer for analysis and processing.[8] This study uses a multi-layer perceptron (MLP) model with a single hidden layer, and the initial learning rate is 0.3. In a random forest (RF) model, the entire random forest is composed of 500 decision trees (ntree = 500), and each decision tree randomly selects 8 variables (mtry = 8) from

40 variables to build a decision tree. Supplementary material 1 illustrates the source code.

Validation

As mentioned before, all data is divided into training set (n=2883) and test set (n=2931). In the training stage,

Table 1 Hyperparameter Tuning in Models

Model	Hyperparameter	Range	Increment	Final Setting
DT	Minimum of number of instances per leaf	2–18	2	2
SVM	Gamma	0.01	0.01	0.02
ANN	Number of hidden neurons	3–12	1	8
RF	Mtry	2–11	1	8
	Number of trees	100–500	100	500

k-fold cross-validation was used to train the model. The training set was divided into 10 parts through non-repetitive sampling. Each time, 9 parts were extracted for model training, and the remaining 1 part was used for model verification. The process was repeated for 10 times, obtaining 10 different models. The test set (n=2931) was used for testing to get the average of the output results as the final indicator (shown in Figure 1). Then, another 50% of the test set (n=2931) was entered to test the performance of the model. The prediction output by the model was compared with the actual diagnosis result to obtain the final result. The prediction performance of models can be significantly influenced by the hyper-parameter setting, such as DT's minimum of number of instances per leaf, SVM's gamma, ANN's number of hidden neurons and RF's mtry and number of trees. The details of the parameter tuning for each model are listed in Table 1.

Model Performance

The Norton scale was used to evaluate the risk of the data in the test group through a two-person team. Specifically, the high risk had a low Norton score (≤15) and an intermediate Norton score (16–18), and the low risk had a high Norton score (≥19).[9,10] The scoring method is based on previous research. The score ≤18 is considered as predictive positive, and the score >18 is considered as predictive negative. The predictive result of Norton scale is calculated following the same performance evaluation method used by the machine learning model. The result is compared with that of the four machine prediction models developed in this research.

The common performance indicators in machine learning including accuracy, recall, precision, F1 value, and ROC curve area (Area Under Curve, AUC) are used in this study. The accuracy of the model was evaluated by calculating the confusion matrix and comparing the ROC curve. Then, we performed model calibration and evaluated the performance of the model after calibration by comparing histograms and reliability diagrams. SPSS V.22.0 software (IBM, Armonk, New York, USA) was used for descriptive statistics.[1] All model analysis was also performed in R language (version 2.9.0 for Windows, http://www.r-project.org).

Results
Basic Features and Potential Predictors

Information of the included patients is shown in Table 1. 1673 patients (28.78%) had pressure ulcer, and their average age was 64.34. The proportion of wards vs ICU was 1504 to 169. Their average length of stay was 8.15 days. 4141 patients (71.22%) did not have pressure ulcer, and their average age was 51.89. The proportion of wards vs ICU was 3557 to 564. The average length of stay was 7.89 days (see Table 2). Multivariate logistic regression analysis indicated that a total of 19 important variables of pressure ulcer risk factors exhibited significant statistical differences. These factors are listed in Table 3.

Performance Evaluation of Four Prediction Models

The test results of the four models are listed in Table 3. The SVM model achieves an accuracy of 94.94%, a recall of 93.90%, a precision of 96.90%, and an F1 value of 94.42%. Compared with the SVM model, the DT model obtains better performance in all indicators. Especially, the F1 value is 3.57% higher than that of the SVM model. The RF model achieves the best performance, with an accuracy of 99.88%, a recall of 99.88%, a precision of 99.93% and an F1 value of 99.88%. The excessive high value of the indicators may be related to the overfitting of this model. The indicators of ANN obtain the lowest value among the four models, showing an accuracy of 79.02%, a recall of 87.21%, a precision of 90.89%, and an F1 value of

Table 2 General Information and Maternal Characteristics of Pressure Ulcer Patients

Variables	Pressure Ulcer Group (n=1673) n/ X̄±s	No Pressure Ulcer Group (n=4141) n/ X̄±s	P
Gender			
Male	1031	2415	0.020
Female	642	1726	
Age (year)	64.34±18.30	51.89±17.55	<0.001
Height (cm)	165.82±13.15	163.54±16.09	0.003
Weight (kg)	65.62±12.32	63.96±14.24	0.014
Total intake (mL)	1577.66±1917.74	1508.15±2221.84	<0.001
Total output (mL)	1786.26±1046.73	2105.70±1058.91	<0.001
Body temperature (°C)	36.78±0.53	36.62±0.64	<0.001
Systolic blood pressure (mmHg)	72.42±10.70	73.12±10.00	<0.001
Blood glucose	9.22±2.44	8.61±2.39	<0.001
Length of stay (days)	8.15±2.87	7.89±3.64	<0.001
Stay in bed	279 (16.68%)	215 (5.19%)	<0.001
Restraint bands	147 (8.79%)	198 (4.78%)	<0.001
Surgery	108 (6.46%)	188 (4.54%)	0.003
Diarrhea	127 (7.59%)	192 (4.64%)	<0.001
Diabetes	72 (4.30%)	113 (2.73%)	0.002
Fracture	252 (15.06%)	57 (1.38%)	<0.001
Norton pressure ulcer assessment			
Total Score	13.78±3.14	18.74±2.38	<0.001
Incontinence score	3.38±0.90	3.96±0.31	<0.001
Activity score	1.74±1.04	3.52±0.90	<0.001
Mind score	3.75±0.70	3.94±0.33	<0.001
Nutritional Assessment	1.33±0.95	0.27±0.57	<0.001
Acceptance of passive turning over	1019 (60.91%)	1282 (30.96%)	<0.001

82.92%. All four machine learning models exhibited an improved performance compared with Norton scale, suggesting that the accuracy of Norton scale for predicting pressure ulcer is worse than that of machine learning model. Indeed, the accuracy and precision of Norton scale are both low. In addition, the AUCs of all four prediction models are higher than 0.95, indicating good fitting effect (see Table 4).

ROC curves of the four models are shown in Figure 3. According to the results, it can be seen that the prediction performance of the four models is quite different, while all the models exhibit acceptable ROC curves and prediction

efficiency. Meanwhile, compared with the ANN model, the other three models show higher predictive accuracy and diagnostic value. Besides, the RF model achieves the best prediction performance among the four models. According to the ROC curves, the four machine learning models achieve higher prediction accuracy than the conventional Norton scale (shown in Figure 3). This may be attributed to the multi-variate nature of machine learning model and its training process. Meanwhile, different machine learning models show consistent accuracies, indicating that the prediction performance of machine learning can be improved by data processing. In addition,

Table 3 Logistic Regression Analysis on Important Variables

Variable	Regression Coefficients	SE(b)	Wald value	P value	OR value	95% CI of OR value	
						Lower Limit	Upper Limit
Constant	−9.560	4.307	4.927	0.026			
Age	0.019	0.004	28.776	<0.001	1.019	1.012	1.027
Weight	0.012	0.005	6.157	0.013	1.012	1.003	1.022
Diarrhea	0.626	0.238	6.921	0.009	1.870	1.173	2.981
Stay in bed	0.933	0.188	24.513	<0.001	2.543	1.757	3.679
Restraint bands	0.381	0.188	4.092	0.043	1.463	1.012	2.116
Surgery	−0.736	0.151	23.632	<0.001	0.479	0.356	0.645
Norton pressure ulcer assessment Total score	2.138	0.308	48.081	<0.001	8.486	4.637	15.532
Norton pressure ulcer assessment Incontinence score	−0.833	0.101	68.063	<0.001	0.435	0.357	0.530
Norton pressure ulcer assessment Activity score	−1.145	0.069	274.010	<0.001	0.318	0.278	0.364
Norton pressure ulcer assessment Mind score	0.403	0.114	12.401	<0.001	1.496	1.196	1.873
Acceptance of passive turning over	−0.504	0.021	579.498	<0.001	0.604	0.580	0.629
Nutritional assessment	0.539	0.079	46.286	<0.001	1.714	1.467	2.002
Total intake	0.577	0.232	6.172	0.013	1.780	1.129	2.805
Total output	−0.595	0.183	10.540	0.001	0.552	0.385	0.790
Body temperature	0.215	0.109	3.927	0.048	1.240	1.002	1.534
Systolic blood pressure	0.011	0.004	6.536	0.011	1.011	1.002	1.019
Blood glucose	0.053	0.017	9.569	0.002	1.054	1.019	1.090
Diabetes	0.435	0.190	5.254	0.022	1.545	1.065	2.241
Fracture	1.250	0.165	57.475	<0.001	3.490	2.526	4.821

Table 4 Comparison of the Prediction Performance of the Four Pressure Ulcer Prediction Models

Model (n=2931)	TP	TN	FP	FN	Accuracy	Recall	Precision	F1 Value	AUC	AUPRC
SVM	770	2.070	41	50	94.94%	93.90%	96.90%	94.42%	0.9940	0.9103
DT	804	2.094	17	16	97.93%	98.05%	98.87%	97.99%	0.9960	0.9607
RF	814	2.075	1	1	99.88%	99.88%	99.93%	99.88%	0.9999	0.9910
ANN	648	2.016	172	95	79.02%	87.21%	90.89%	82.92%	0.9590	0.8542
Norton scale	437	1261	490	850	47.14%	33.95%	55.89%	39.48%	0.5709	0.8339

Abbreviations: TP, true positive; FN, false negative; TN, true negative; FP, false positive; AUC, area under the curve; AUPRC, area under the precision recall curve.
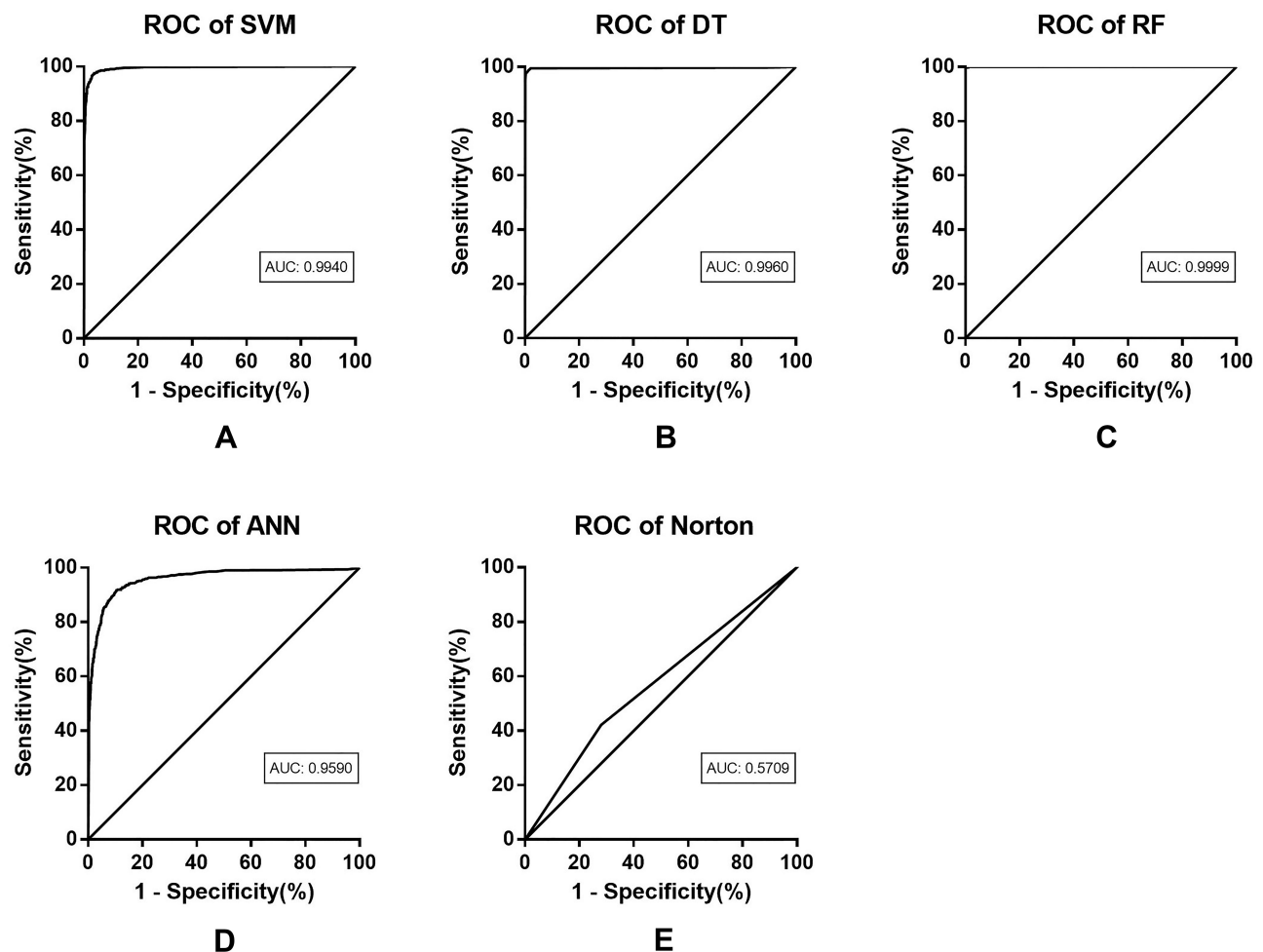
Figure 3 Performance metrics of pressure ulcer prediction models on the test data set. Based on the prediction results of the model, the ROC curves are drawn, including SVM (**A**), DT (**B**), RF (**C**), and ANN (**D**). The results are compared with the ROC curve of Norton scale (**E**). Norton scale is inferior to the machine learning model in terms of ROC curve or AUC value. Graphically, DT and RF achieve similar performance, but RF obtains higher prediction accuracy in terms of AUC value.

traditional evaluations are inferior to the algorithms based on feature values, and the reason may be that the scale was developed earlier and did not incorporate multiple variables. The machine learning technology has advantages in adjustable algorithm and real-time data.

The result of model calibration is shown in Figure 4. The reliability diagrams of ANN, DT, SVM, and RF all show an S-shape, which are significantly improved after calibration. After calibration, the predicted value of output tends to be distributed in two levels of 0 and 1. Model calibration can significantly improve the performance of pressure ulcer prediction models. Moreover, the RF model still has the best performance after calibration. Model calibration plays an important role in improving the performance of the model, which will make the model more useful in clinical

applications, especially as a clinical decision support tool for ulcer risk scoring.

Discussion

This study presents a prediction model constructed based on retrospective data. The characteristics of high-risk factors within 48 hours before the occurrence of pressure ulcer are taken into consideration. After weighting these high-risk variables, the model training is constructed and implemented, and an optimal model is selected. The output result gives the probability of pressure ulcer occurring within 48 hours of the data time point (shown in Figure 1). Therefore, when the model captures the data of high-risk variables, it can predict whether the target patient will suffer from pressure ulcer in the next 48 hours.
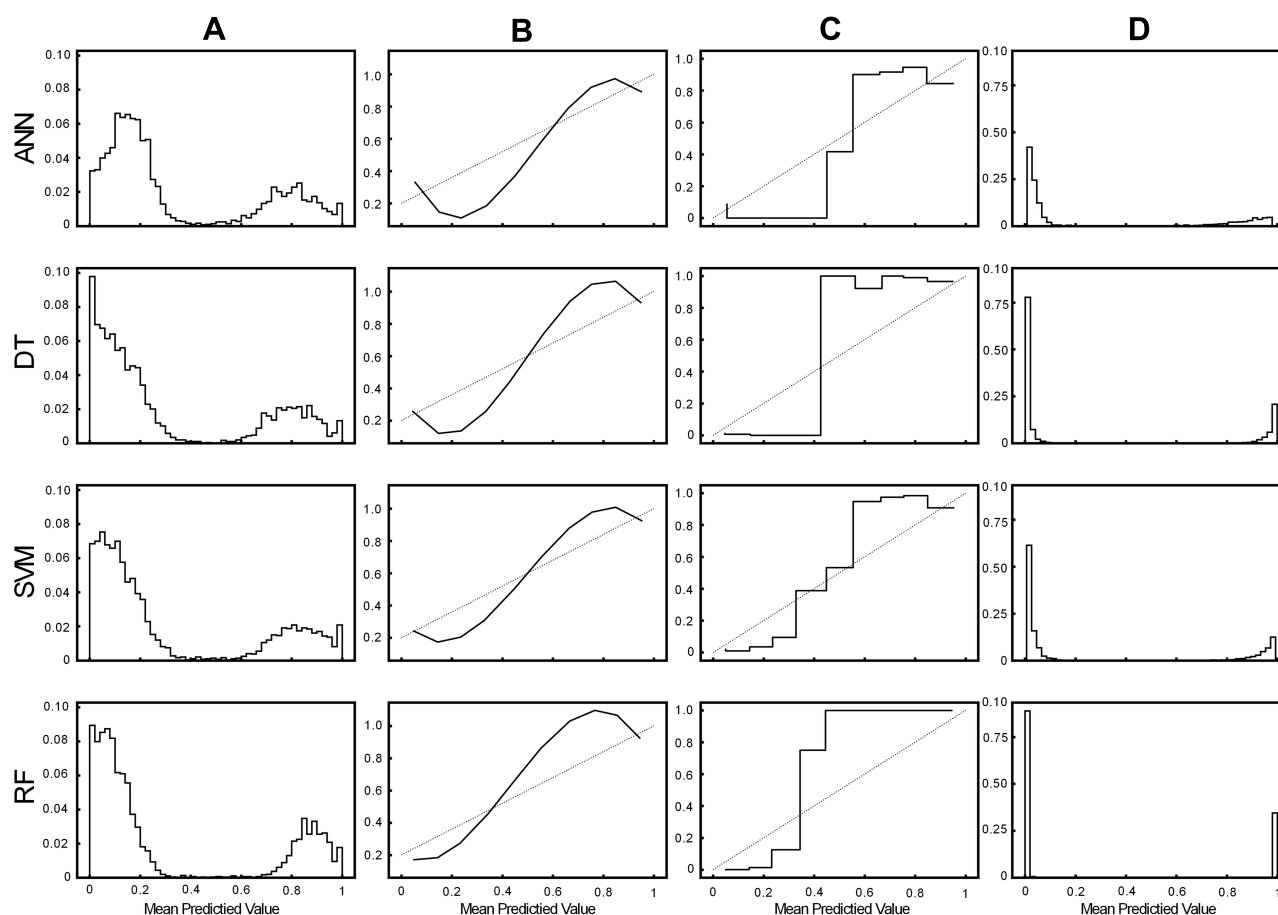
Figure 4 Histograms and reliability diagrams for models. (**A–C**) are the histograms and reliability diagrams before model calibration. The column (**D**) is the histogram of the model after calibration.

The related variables related to the occurrence of pressure ulcer in patients are found in this study, including age, weight, total intake, total output, body temperature, systolic blood pressure, blood glucose, diarrhea, stay in bed, restraint bands, surgery, total score of pressure ulcer assessment, acceptance of passive turning over, nutritional assessment, diabetes, and fracture. Several key variables such as age, weight, total intake, total output, body temperature, systolic blood pressure and blood glucose are consistent with those reported in other studies.[11–14] Systolic blood pressure is highly correlated with pressure ulcers, which may be related to the patient's hemodynamic factors.[15] Diarrhea is an important factor, and it is closely related to pressure ulcer, especially the one occurred around anus.[16] It promotes flushing, edema, and pain of the perianal skin, and can even cause skin ulcers and infections. Bedridden patients are easy to suffer from localized damage to the skin and subcutaneous tissues.[17] This study also found a correlation between bed rest and pressure ulcer. According to a multicenter survey, 31.4%

of long-term bedridden patients suffer from pressure ulcer among hospitalized patients.[18] The incidence of pressure ulcer in bedridden patients ranges from 13.3% to 57.6%.[19–23] The average time of stay in surgery is about 8.6 days (slightly less than the hospital level of 9.1 days in China), which is consistent with the researches.[24–26] This study found that surgery is related to pressure ulcer, and Aloweni et al confirmed this result. Other studies have confirmed that the limitation and irregular use of the restraint belt are related to pressure ulcer.[27]

The analysis of acceptance of passive turning over, pressure ulcer evaluation, and nutritional evaluation helps us predict the occurrence of pressure ulcer from the perspective of prevention. Conventional scales such as Norton, Breslow, and Waterlow have all evaluated the nutritional status. However, patients receiving treatment in the hospital will receive nutritional intervention treatment and some nursing measures to prevent pressure ulcer, such as turning over.[28] The use of these measures will affect the occurrence of pressure ulcer, making it

necessary to consider these positive and related measures for predicting pressure ulcer. The advantages of machine learning are related to the database, which makes the response time of variable analysis very fast. Based on the important variables investigated by this research, nurses can save the link of scale evaluation during the pressure ulcer evaluation, which contributes to a reduced evaluation time and more accurate result for pressure ulcer prediction. The selection of important variables related to pressure ulcer is crucial, because it directly affects the accuracy of the feature extraction and output results of the machine learning model. More variables related to pressure ulcer can be exploited to upgrade the model in future research.

Meanwhile, the pressure ulcer prediction models constructed based on these 19 factors have achieved satisfactory prediction performance, and the RF model performs the best. The machine learning model shows a comprehensive predictive ability for pressure ulcer, a complex multi-factor disease. Compared with Norton scale, the proposed machine learning model exhibited an improved prediction performance. It may be that the individual differences of pressure ulcer-related variables cannot reflected by the assessment scales. However, these problems can be handled by machine learning, through obtaining more variable data, performing multi-variable tasks, screening and processing the multi-factor characteristics of pressure ulcer, and outputting prediction results highly correlated with pressure ulcer. Additionally, machine learning adopts more optimized algorithms and more explicit feature extraction, which leads to more optimized data fitting than traditional tools.

In the process of constructing the machine learning prediction model in this study, clarifying the related factors of pressure ulcer is helpful for the model training and make the parameters of the pressure ulcer prediction model more accurate. Three methods are used in this study to improve the prediction performance of the model, including controlling the number of variables, improving the test process, and selecting the appropriate model.[29]

Logistic regression is exploited to refine the characteristics of risk, reduce low-related variables, and find enough important variables. The low prediction accuracy of the prediction model may be related to the less included factors.[30] It is impossible to accurately determine whether the patient is at risk of nursing adverse events with limited factors. But more variables will make the generated machine learning model over-fitted. In this case, a preliminary selection of predictive variables is conducted based on the clinical knowledge. In stage 2 of the model construction, a logistic regression model

was used to determine the variables (variables with $P<0.05$ were selected), so that the number of variables was limited. These variables have a high correlation with pressure ulcer, and the prediction accuracy of the machine learning model is greatly improved. As a current mainstream method, K-fold cross-validation is used in the train set to improve the efficiency of model training.[31] For small sample data, the model performance cannot be improved by only setting train and test data sets. The nested use of K-fold cross-validation in the train set can achieve the expected prediction effect.[32] Also, the verification in the train set can maximize the use of data to a certain extent, for the amount of pressure ulcer data in this study is not very huge.

The comparison between models is also an important way to improve model's prediction performance. Different machine learning algorithms are suitable for different clinical problems, and pressure ulcer prediction on digital medical record data is often regarded as a linear fitting problem or classification problem. Previous studies usually selected the best model by comparing the performance of multiple models.[33,34] Because the machine learning model has different degrees of adaptation when processing different data types. The most common machine learning models are compared to find the subtle differences in the models. Through evaluations of the four models for predicting pressure ulcer adverse event, it was found that the accuracy and recall of the four models were high. RF model showed superior performance than the traditional prediction models. This study adopted a similar method for constructing the pressure ulcer model and obtained similar findings.[35,36] Alderden et al also demonstrated that the machine learning model achieved high accuracy for predicting pressure ulcer, and random forest model obtained the best performance (AUC=0.79).[37] However, more factors are considered in this study, which can better analyze the risk of pressure ulcer. The random forest model proposed by Hu et al provided a good inspiration for pressure ulcer prediction. In this study, the average precision of DT, logistic regression and RF is 0.969, 0.799, and 0.998, respectively.[38] It indicates that the random forest algorithm is more efficient in processing classification problems. In addition, other studies use decision tree models to predict pressure ulcer.[39] However, DT is not as good as RF in fitting. The low predictive ability of the ANN model shows that there may be an "overfitting" phenomenon. Also, the 19 variables may bring a large computational burden for the ANN model. The performance of ANN on categorical variables is not as good as that of other models. Therefore, not all machine learning models are suitable for

pressure ulcer prediction, suggesting that researchers need to be more cautious in choosing machine learning models for making disease predictions in the future.

Limitations of the Study

At the time of writing this paper, the electronic medical system contains text description rather than image of the pressure ulcer of the patient, and the description content could not be unified. In this case, it can be determined whether the patient has a pressure ulcer, but the classification and prognosis of pressure ulcer are hard. In the future, the description of pressure ulcer event needs to be standardized in the electronic form. At the hospital where the study is located, the incidence of pressure ulcer is only 2.5%, while the average incidence reported is 0.4%～38%. Due to heavy workload, nurses may neglect the care for patients, especially some patients with mild pressure ulcer or patients about to be discharged. This reduces the model's ability to identify mild pressure ulcers, and this issue remains to be confirmed. Therefore, single-center research may make the processing of diverse data a challenge, requiring more case data from multiple centers. This challenge could further limit the initial clinical application of the proposed model. Additionally, there is no consensus on the risks of pressure ulcer so far. Although many factors are included in this study, the risks taken into consideration for constructing the pressure ulcer prediction model cannot cover all potential pressure ulcer risks. Limited to the hospital digital medical record database, only a strong correlation exists between the model's inference and the characteristics, but it is still impractical to draw a complete causal relationship. Finally, it is essential to build a model based on a new pressure ulcer medical record database that covers all factors in the future, and the use of multi-center data may significantly promote the clinical application of the model constructed in this study.

Conclusion

Four machine learning models for predicting the pressure ulcer adverse events are studied in this paper. Experimental results based on the common evaluation indictors show that the four models achieve high capabilities in predicting pressure ulcer adverse events. Also, the horizontal comparisons of the models show that the random forest and the decision tree model have better prediction performance and are more suitable for predicting pressure ulcer adverse events. The construction of the pressure ulcer adverse event prediction model provides new early warning tools for pressure ulcer risk in the clinic and improves the feasibility of personalized care. Although the

machine learning has been widely used to predict many diseases, there are few cases for pressure ulcer prediction. The pressure ulcer prediction model based on machine learning can be used as a pressure ulcer prevention tool with broad prospects.

Author Contributions

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; Song Jie, Gao Yuan, and Pi Hongying took part in drafting the article or revising it critically for important intellectual content; All authors agreed to submit to the current journal, gave final approval of the version to be published and agree to be accountable for all aspects of the work.

Disclosure

The authors have no conflicts of interest to disclose.

References

1. Wound O; Continence Nurses Society-Wound Guidelines Task F. WOCN 2016 guideline for prevention and management of pressure injuries (ulcers): an executive summary. *J Wound Ostomy Contin Nurs*. 2017;44(3):241–246. doi:10.1097/WON.0000000000000321
2. Tomova-Simitchieva T, Akdeniz M, Blume-Peytavi U, et al. The epidemiology of pressure ulcer in Germany: systematic review. *Gesundheitswesen*. 2019;81(6):505–512. doi:10.1055/s-0043-122069
3. Konya C, Sanada H, Sugama J, et al. Skin injuries caused by medical adhesive tape in older people and associated factors. *J Clin Nurs*. 2010;19(9–10):1236–1242. doi:10.1111/j.1365-2702.2009.03168.x
4. Mervis JS, Phillips TJ. Pressure ulcers: pathophysiology, epidemiology, risk factors, and presentation. *J Am Acad Dermatol*. 2019;81 (4):881–890. doi:10.1016/j.jaad.2018.12.069
5. Jhee JH, Lee S, Park Y, et al. Prediction model development of late-onset preeclampsia using machine learning-based methods. *PLoS One*. 2019;14(8):e0221202. doi:10.1371/journal.pone.0221202
6. Moore Z, Cowman S, Posnett J. An economic analysis of repositioning for the prevention of pressure ulcers. *J Clin Nurs*. 2013;22(15–16):2354–2360. doi:10.1111/j.1365-2702.2012.04310.x
7. Tingle J. World Health Organization: providing global leadership for patient safety. *Br J Nurs*. 2017;26(13):778–779. doi:10.12968/bjon.2017.26.13.778
8. Benin AL, Fodeh SJ, Lee K, et al. Electronic approaches to making sense of the text in the adverse event reporting system. *J Healthc Risk Manag*. 2016;36(2):10–20. doi:10.1002/jhrm.21237
9. Rabinovitz E, Finkelstein A, Ben Assa E, et al. Norton scale for predicting prognosis in elderly patients undergoing trans-catheter aortic valve implantation: a historical prospective study. *J Cardiol*. 2016;67(6):519–525. doi:10.1016/j.jjcc.2016.01.017
10. Silber H, Shiyovich A, Gilutz H, et al. Decreased Norton's functional score is an independent long-term prognostic marker in hospital survivors of acute myocardial infarction. Soroka Acute Myocardial Infarction II (SAMI-II) project. *Int J Cardiol*. 2017;228:694–699. doi:10.1016/j.ijcard.2016.11.112
11. Waterlow J. Tissue viability. Prevention is cheaper than cure. *Nurs Times*. 1988;84(25):69–70.
12. Maklebust J. Pressure ulcers: etiology and prevention. *Nurs Clin North Am*. 1987;22(2):359–377.

13. Lahmann NA, Kottner J. Relation between pressure, friction and pressure ulcer categories: a secondary data analysis of hospital patients using CHAID methods. *Int J Nurs Stud*. 2011;48 (12):1487–1494. doi:10.1016/j.ijnurstu.2011.07.004

14. Okuwa M, Sanada H, Sugama J, et al. A prospective cohort study of lower-extremity pressure ulcer risk among bedfast older adults. *Adv Skin Wound Care*. 2006;19(7):391–397. doi:10.1097/00129334-200609000-00017

15. Brindle CT, Malhotra R, O'rourke S, et al. Turning and repositioning the critically ill patient with hemodynamic instability: a literature review and consensus recommendations. *J Wound Ostomy Continence Nurs*. 2013;40(3):254–267. doi:10.1097/WON.0b013e318290448f

16. Benoit RA, Watts C. The effect of a pressure ulcer prevention program and the bowel management system in reducing pressure ulcer prevalence in an ICU setting. *J Wound Ostomy Continence Nurs*. 2007;34 (2):163–177. doi:10.1097/01.WON.0000264830.26355.64

17. Edsberg LE, Black JM, Goldberg M, et al. Revised national pressure ulcer advisory panel pressure injury staging system: revised pressure injury staging system. *J Wound Ostomy Contin Nurs*. 2016;43 (6):585–597. doi:10.1097/WON.0000000000000281

18. Brsting TE, Tvedt CR, Skogestad IJ, et al. Prevalence of pressure ulcer and associated risk factors in middle- and older-aged medical inpatients in Norway. *J Clin Nurs*. 2018;27(3–4):e535–e543. doi:10.1111/jocn.14088

19. Briggs M, Collinson M, Wilson L, et al. The prevalence of pain at pressure areas and pressure ulcers in hospitalised patients. *BMC Nurs*. 2013;12. doi:10.1186/1472-6955-12-19

20. Gunningberg L, Hommel A, Baath C, et al. The first national pressure ulcer prevalence survey in county council and municipality settings in Sweden. *J Eval Clin Pract*. 2013;19(5):862–867. doi:10.1111/j.1365-2753.2012.01865.x

21. Vocci MC, Fontes CMB, Abbade LPF. Pressure injury in the pediatric population: cohort study using the Braden Q scale. *Adv Skin Wound Care*. 2018;31(10):456–461. doi:10.1097/01.ASW.0000542529.94557.0a

22. Silva MS, Garcia TR. Risk factors for pressure ulcer in bedridden patients. *Rev Bras Enferm*. 1998;51(4):615. doi:10.1590/S0034-71671998000400007

23. Bianchetti A, Zanetti O, Rozzini R, et al. Risk factors for the development of pressure sores in hospitalized elderly patients: results of a prospective study. *Arch Gerontol Geriatr*. 1993;16(3):225. doi:10.1016/0167-4943(93)90034-F

24. Scheib SA, Thomassee M, Kenner JL. Enhanced Recovery After Surgery (ERAS) in gynecology: a review of the literature. *J Minim Invasive Gynecol*. 2019;26(2):327–343. doi:10.1016/j.jmig.2018.12.010

25. Ying S, Qiu L, Ping S. Analysis of influencing factors and countermeasures of average hospital stay in surgery. *Modern Hosp Manag*. 2015;3:11–13.

26. National Health and Family Planning Commission of the People's Republic of China. National medical services from January to October 2017. Available from: http://www.nhfpc.gov.cn/mohwsbwstjxxzx/s7967/201712/9e82bc727b8d4c6b. Accessed December 29, 2017.

27. Aloweni F, Ang SY, Fook-Chong S, et al. A prediction tool for hospital-acquired pressure ulcers among surgical patients: surgical pressure ulcer risk score. *Int Wound J*. 2019;16(1):164–175. doi:10.1111/iwj.13007

28. Fulbrook P, Anderson A. Pressure injury risk assessment in intensive care: comparison of inter-rater reliability of the COMHON (Conscious level, Mobility, Haemodynamics, Oxygenation, Nutrition) Index with three scales. *J Adv Nurs*. 2016;72(3):680–692. doi:10.1111/jan.12825

29. Li X, Zhang X, Zhu J, et al. Depression recognition using machine learning methods with different feature generation strategies. *Artif Intell Med*. 2019;99:101696. doi:10.1016/j.artmed.2019.07.004

30. Xiong Y, Wang Z, Jiang D. A fine-grained Chinese word segmentation and part-of-speech tagging corpus for clinical text. *BMC Med Inform Decis Mak*. 2019;19(Suppl2):66. doi:10.1186/s12911-019-0770-7

31. Lopez-Del Rio A, Nonell-Canals A, Vidal D, et al. Evaluation of cross-validation strategies in sequence-based binding prediction using deep learning. *J Chem Inf Model*. 2019;59(4):1645–1657. doi:10.1021/acs.jcim.8b00663

32. Vabalas A, Gowen E, Poliakoff E, et al. Machine learning algorithm validation with a limited sample size. *PLoS One*. 2019;14(11): e0224365. doi:10.1371/journal.pone.0224365

33. Struck AF, Rodriguez-Ruiz AA, Osman G, et al. Comparison of machine learning models for seizure prediction in hospitalized patients. *Ann Clin Transl Neurol*. 2019;6(7):1239–1247. doi:10.1002/acn3.50817

34. Choi BG, Rha SW, Kim SW, et al. Machine learning for the prediction of new-onset diabetes mellitus during 5-year follow-up in non-diabetic patients with cardiovascular risks. *Yonsei Med J*. 2019;60(2):191–199. doi:10.3349/ymj.2019.60.2.191

35. Liu Y, Ye S, Xiao X, et al. Machine learning for tuning, selection, and ensemble of multiple risk scores for predicting type 2 diabetes. *Risk Manag Healthc Policy*. 2019;12:189–198. doi:10.2147/RMHP.S225762

36. Nindrea RD, Aryandono T, Lazuardi L, et al. Diagnostic accuracy of different machine learning algorithms for breast cancer risk calculation: a meta-analysis. *Asian Pac J Cancer Prev*. 2018;19 (7):1747–1752. doi:10.22034/APJCP.2018.19.7.1747

37. Alderden J, Pepper GA, Wilson A, et al. Predicting pressure injury in critical care patients: a machine-learning model. *Am J Crit Care*. 2018;27(6):461–468. doi:10.4037/ajcc2018525

38. Hu YH, Lee YL, Kang MF, et al. Constructing inpatient pressure injury prediction models using machine learning techniques. *Comput Inform Nurs*. 2020;38(8):415–423. doi:10.1097/CIN.0000000000000604

39. Moon M, Lee SK. Applying of decision tree analysis to risk factors associated with pressure ulcers in long-term care facilities. *Healthc Inform Res*. 2017;23(1):43–52. doi:10.4258/hir.2017.23.1.43